

Exactness of OPF Relaxation on Three-phase Radial Networks with Delta Connections

Fengyu Zhou, *Student Member, IEEE*, Ahmed S. Zamzam, *Member, IEEE*, Steven H. Low, *Fellow, IEEE*, and Nicholas D. Sidiropoulos, *Fellow, IEEE*

Abstract—Simulations have shown that while semi-definite relaxations of AC optimal power flow (AC-OPF) on three-phase radial networks with only wye connections tend to be exact, the presence of delta connections seem to render them inexact. This paper shows that such inexactness originates from the non-uniqueness of relaxation solutions and numerical errors amplified by the non-uniqueness. This finding motivates two algorithms to recover the exact solution of AC-OPF in unbalanced distribution networks featuring both wye and delta connections. In simulations using IEEE 13, 37 and 123-bus systems, the proposed algorithms provide exact optimal solutions up to numerical precision.

Index Terms—Optimal power flow, semi-definite relaxation, distribution networks, delta-connected devices.

I. INTRODUCTION

OPTIMAL power flow (OPF) is a mathematical program that finds an optimal operating point for a power grid subject to laws of physics and operational constraints [1]. OPF formulated under the AC model is known to be both nonconvex and NP-hard to solve [2], [3]. Commonly used methods to approximately solve this non-convex problem include, e.g., the Newton-Raphson method [4] and various linearizations [5], [6], [7]. Another approach that has emerged over the last decade or so is to relax the OPF problem to a convex program, such as semi-definite program (SDP) relaxations and second-order cone program (SOCP) relaxations [8], [9]. For single-phase radial networks (i.e., networks with a tree topology) as well as the single-phase equivalent of a balanced three-phase radial network, simulations have shown that these relaxations often yield solutions that are also global optima of the original nonconvex problems [10]. Sufficient conditions that guarantee

exact relaxations for single-phase radial networks have subsequently been proved. There is now sizeable literature on OPF relaxations; see, e.g., comprehensive surveys in [10], [11] for pointers to various contributions including many earlier surveys on OPF. In [12], the authors provided a comprehensive comparison between nonlinear methods and relaxation for test cases of different sizes. Many recent works further refine and improve the convexification of OPF problems and exhibit promising performance [13], [14].

Most of this literature focuses on single-phase models, but distribution systems have multiple phases that are increasingly unbalanced as distributed energy resources continue to grow. Mathematically, we can identify each bus-phase pair of a three-phase radial network with an equivalent single-phase bus and transform it into a single-phase equivalent circuit with a meshed topology [15], [16]. Most analytical results on exact relaxation for meshed networks are restricted to weakly-cyclic networks [17], but the single-phase equivalent of a three-phase radial network is beyond this class. SDP relaxation has been generalized in [18], [19] to three-phase radial networks with only wye connections and shown to be exact in the simulation of several test cases. A recent work [20] proves a sufficient condition for exact relaxation in this case. Besides SDP relaxation, [21] also provides an iterative algorithm for three-phase networks, also without delta connections.

Semi-definite relaxation is recently extended in [22] to networks with both wye and delta connected devices by introducing a new positive semi-definite matrix that represents the outer product of voltages and phase-to-phase currents in the delta connections (matrix $\mathbf{M}^{\mathbf{v}, \mathbf{x}, \rho}(j)$ in (8b) below). Simulation results in [22] showed that, surprisingly, this matrix was never rank-1 at an optimal solution of the relaxation. This seems to suggest that the SDP relaxation was inexact in these simulations. In this paper, we show that even though the matrix $\mathbf{M}^{\mathbf{v}, \mathbf{x}, \rho}(j)$ fails to attain rank 1, an exact solution can still be recovered under certain conditions; see Theorem 2 and Remark 1. The inexactness in previous works is due to two issues. First, optimal solutions to the SDP relaxation in these simulations are generally not unique, and the exact solution is only one of them which is not returned by the solver. Second, such non-uniqueness could significantly amplify the numerical error and make it computationally challenging to recover the exact solution. We propose two variants of the standard SDP relaxation that address both issues. The first algorithm post-processes the relaxation solution and tends to provide lower cost but larger constraint violation, while the second algorithm adds a penalty term to the cost and tends to provide higher

Fengyu Zhou and Steven H. Low are with the Department of Electrical Engineering, California Institute of Technology, Pasadena, CA 91125 USA (e-mail: {f.zhou, slow}@caltech.edu). Ahmed S. Zamzam is with the National Renewable Energy Laboratory, Golden, CO 80401 USA (e-mail: ahmed.zamzam@nrel.gov). Nicholas D. Sidiropoulos is with the Electrical and Computer Engineering Department, University of Virginia, Charlottesville, VA 22901 USA (email: nikos@virginia.edu).

This work was funded in part by NSF through grants CCF 1637598 and ECCS 1931662. This work was authored in part by the National Renewable Energy Laboratory, operated by Alliance for Sustainable Energy, LLC, for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308. The work of A. S. Zamzam was supported by the Laboratory Directed Research and Development (LDRD) Program at the National Renewable Energy Laboratory. The views expressed in the article do not necessarily represent the views of the DOE or the U.S. Government. The U.S. Government retains and the publisher, by accepting the article for publication, acknowledges that the U.S. Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this work, or allow others to do so, for U.S. Government purposes.

cost but smaller constraint violation. Simulations of both algorithms corroborate the theoretical results and show that they can recover exact solutions for three IEEE distribution feeders.

To summarize the main contribution, this paper first explains why conventional semi-definite relaxation is often inexact when delta connections are present. Then we propose two algorithms and prove that they can recover exact solutions under certain conditions. As a byproduct, we also prove that two models, bus injection model and branch flow model, are equivalent when we relax the problem.

The remainder of the paper is organized as follows. In Section II, we define the network structure and formulate the three-phase OPF problem in both the bus injection model (BIM) and the branch flow model (BFM). Section III proves that the global optimal solution to the nonconvex OPF problem can be recovered from its relaxation under certain conditions, and two algorithms are presented. Section IV shows the equivalence between BIM and BFM. Finally, in Section V, we apply our algorithms to IEEE 13-, 37- and 123-bus systems.

II. SYSTEM MODEL

A. Network Structure

We study the model proposed in [19], [22]. Let the directed graph representing the electrical network be $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} = \{0, 1, \dots, n\}$ denotes the set of buses, and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ denotes the set of edges; and let $N := |\mathcal{V}| = n + 1$. In this paper, we focus on the case where \mathcal{G} represents a radial network (i.e., a tree) because most distribution networks have a tree topology. Throughout the paper, we will use (graph, vertex, edge) and (power network, bus, line) interchangeably. Without loss of generality, we let bus 0 be the substation bus where the distribution feeder is connected to a transmission network. Suppose the substation also serves as the slack bus, so the voltages at the substation bus are fixed and specified. We use $j \rightarrow k$ to denote a directed edge from bus j to k . In many situations, when we do not care about the direction of the edge, we simply use (j, k) and $j \sim k$ interchangeably to denote an edge connecting bus j and k . That means either $j \rightarrow k$ or $k \rightarrow j$ is in \mathcal{E} . Consider a three-phase line (j, k) characterized by the series impedance matrix $z_{jk} \in \mathbb{C}^{3 \times 3}$. When line (j, k) has three phases, the inverse of z_{jk} , denoted as y_{jk} , is the admittance of line (j, k) . If branch (j, k) has less than three phases, then we fill the rows and columns of z_{jk} corresponding to the missing phases with zeros, and we let the admittance matrix y_{jk} be the pseudo-inverse of z_{jk} . Last, let $y_j \in \mathbb{C}^{3 \times 3}$ denote the admittance of a shunt device connected to bus j .¹

For each bus j , let the voltages of all three phases at bus j be collected in the vector $\mathbf{V}_j \in \mathbb{C}^3$. We use \mathbf{V}_j^ϕ for $\phi \in \{a, b, c\}$ to indicate the voltage of phase ϕ . The voltage \mathbf{V}_0 at slack bus 0 is known and denoted by \mathbf{V}_{ref} . Let $\mathbf{V} = [\mathbf{V}_0^\top, \mathbf{V}_1^\top, \dots, \mathbf{V}_n^\top]^\top$ collect the voltages for the entire network. Similarly, we use \mathbf{s}_j^ϕ to denote the bus injection for phase ϕ at bus j ; and we

denote \mathbf{s}_j and \mathbf{s} as the injections at bus j and in the entire network, respectively.

For delta connected components, we use $\mathbf{I}_{\Delta,j} \in \mathbb{C}^3$ to collect the delta line currents for phases in $\{ab, bc, ca\}$. Define

$$\Gamma := \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ -1 & 0 & 1 \end{bmatrix}.$$

Therefore, the complex power injections of the delta connected components at bus j can be expressed as $\mathbf{s}_{\Delta,j} = \text{diag}(\Gamma \mathbf{V}_j \mathbf{I}_{\Delta,j}^H)$. The net nodal injections contributed by delta connections at bus j are given by $-\text{diag}(\mathbf{V}_j \mathbf{I}_{\Delta,j}^H \Gamma)$. Assume that the operation regions for \mathbf{s}_j and $\mathbf{s}_{\Delta,j}$ at bus j are convex compact sets \mathcal{S}_j and $\mathcal{S}_{\Delta,j}$, respectively.

The AC power flow equations are:

$$\begin{aligned} & \mathbf{s}_j - \text{diag}(\mathbf{V}_j \mathbf{I}_{\Delta,j}^H \Gamma) - \text{diag}(\mathbf{V}_j \mathbf{V}_j^H y_j^H) \\ &= \sum_{k:j \sim k} \text{diag}((\mathbf{V}_j \mathbf{V}_j^H - \mathbf{V}_j \mathbf{V}_k^H) y_{jk}^H) \end{aligned} \quad (1a)$$

$$\mathbf{s}_{\Delta,j} = \text{diag}(\Gamma \mathbf{V}_j \mathbf{I}_{\Delta,j}^H). \quad (1b)$$

where (1a) is the power balance equation at bus j and (1b) defines the power through delta connected components.

Similar to [22], we will adopt $\mathbf{X}_j, \rho_j \in \mathbb{C}^{3 \times 3}$ as auxiliary matrices to model the outer products of voltages and currents.

$$\mathbf{X}_j = \mathbf{V}_j \mathbf{I}_{\Delta,j}^H \quad (2a)$$

$$\rho_j = \mathbf{I}_{\Delta,j} \mathbf{I}_{\Delta,j}^H. \quad (2b)$$

We consider an OPF problem that minimizes a continuous convex cost function $f(\mathbf{s}, \mathbf{s}_{\Delta})$ over variables $(\mathbf{s}, \mathbf{s}_{\Delta}, \mathbf{V}, \mathbf{I}_{\Delta})$ subject to power flow equations (1) as well as voltage and injection limits:

$$\begin{aligned} & \underset{\mathbf{s}, \mathbf{s}_{\Delta}, \mathbf{V}, \mathbf{I}_{\Delta}}{\text{minimize}} && f(\mathbf{s}, \mathbf{s}_{\Delta}) \end{aligned} \quad (3a)$$

$$\text{subject to} \quad (1) \quad (3b)$$

$$\mathbf{V}_0 = \mathbf{V}_{\text{ref}} \quad (3c)$$

$$\mathbf{s}_j \in \mathcal{S}_j, \mathbf{s}_{\Delta,j} \in \mathcal{S}_{\Delta,j}, \text{ for } j \in \mathcal{V} \quad (3d)$$

$$\underline{\mathbf{V}} \leq |\mathbf{V}| \leq \bar{\mathbf{V}}. \quad (3e)$$

In (3e), $|\mathbf{V}|$ stands for the modulus of \mathbf{V} elementwise, and $\underline{\mathbf{V}}, \bar{\mathbf{V}} \in \mathbb{R}^{3N}$ are the lower and upper limits for voltage magnitudes, respectively. If the limits are homogeneous across all buses and phases, we can denote them as $\underline{V}\mathbf{1}, \bar{V}\mathbf{1}$, where \underline{V}, \bar{V} are scalars and $\mathbf{1}$ is the all-one vector. Next, we will present power flow equations for three-phase radial networks in both bus injection model (5) and branch flow model (9).

B. Bus Injection Model

The *bus injection model* (BIM) is defined in terms of $(\mathbf{s}, \mathbf{s}_{\Delta}, \mathbf{W}, \mathbf{X}, \rho)$, where we use $\mathbf{W} \in \mathbb{C}^{3N \times 3N}$ to replace $\mathbf{V}\mathbf{V}^H$ in (1). The matrix $\mathbf{W}_{jk} \in \mathbb{C}^{3 \times 3}$ is the (j, k) submatrix of \mathbf{W} . For notational simplicity, we let

$$\mathbf{M}^{\mathbf{W}, \mathbf{X}, \rho}(j) := \begin{bmatrix} \mathbf{W}_{jj} & \mathbf{X}_j \\ \mathbf{X}_j^H & \rho_j \end{bmatrix} \quad \text{for } j \in \mathcal{V}. \quad (4)$$

¹The shunt here refers to a capacitive device at bus j and not the line charging in the Π circuit model.

The power flow model is represented as

$$\begin{aligned} & \mathbf{s}_j - \text{diag}(\mathbf{X}_j \Gamma) - \text{diag}(\mathbf{W}_{jj} \mathbf{y}_j^H) \\ &= \sum_{k:j \sim k} \text{diag}((\mathbf{W}_{jj} - \mathbf{W}_{jk}) \mathbf{y}_{jk}^H) \end{aligned} \quad (5a)$$

$$\mathbf{s}_{\Delta,j} = \text{diag}(\Gamma \mathbf{X}_j) \quad (5b)$$

$$\mathbf{W}_{00} = \mathbf{V}_{\text{ref}} \mathbf{V}_{\text{ref}}^H \quad (5c)$$

$$\mathbf{W} \succeq 0 \quad (5d)$$

$$\text{rank}(\mathbf{W}) = 1 \quad (5e)$$

$$\mathbf{M}^{\mathbf{W}, \mathbf{X}, \rho}(j) \succeq 0 \quad (5f)$$

$$\text{rank}(\mathbf{M}^{\mathbf{W}, \mathbf{X}, \rho}(j)) = 1. \quad (5g)$$

Equations (5f), (5g) are derived from (2) to model the current and power flow of delta connections.

Hence, the AC-OPF problem in BIM formulation is:

$$\begin{aligned} & \text{minimize}_{\mathbf{s}, \mathbf{s}_{\Delta}, \mathbf{W}, \mathbf{X}, \rho} f(\mathbf{s}, \mathbf{s}_{\Delta}) \end{aligned} \quad (6a)$$

$$\text{subject to} \quad (5), (3d). \quad (6b)$$

$$\text{diag}(\mathbf{V} \mathbf{V}^H) \leq \text{diag}(\mathbf{W}) \leq \text{diag}(\overline{\mathbf{V} \mathbf{V}^H}) \quad (6c)$$

C. Branch Flow Model

In a *branch flow model* (BFM), we introduce \mathbf{S} , \mathbf{v} , and ℓ to model the branch power flow, squared voltages, and squared currents, respectively. We let $\mathbf{I}_{jk} := y_{jk}(\mathbf{V}_j - \mathbf{V}_k)$ be the sending-end current from bus j to bus k , and \mathbf{S}_{jk} be the sending-end branch power from j to k . The matrices \mathbf{S} , \mathbf{v} , and ℓ can be written as

$$\mathbf{S} = (\mathbf{S}_{jk} \in \mathbb{C}^{3 \times 3})_{(j \rightarrow k) \in \mathcal{E}}, \quad \mathbf{S}_{jk} = \mathbf{V}_j \mathbf{I}_{jk}^H \quad (7a)$$

$$\mathbf{v} = (\mathbf{v}_j \in \mathbb{C}^{3 \times 3})_{j \in \mathcal{V}}, \quad \mathbf{v}_j = \mathbf{V}_j \mathbf{V}_j^H \quad (7b)$$

$$\ell = (\ell_{jk} \in \mathbb{C}^{3 \times 3})_{(j \rightarrow k) \in \mathcal{E}}, \quad \ell_{jk} = \mathbf{I}_{jk} \mathbf{I}_{jk}^H. \quad (7c)$$

Let

$$\mathbf{M}^{\mathbf{v}, \mathbf{S}, \ell}(j, k) := \begin{bmatrix} \mathbf{v}_j & \mathbf{S}_{jk} \\ \mathbf{S}_{jk}^H & \ell_{jk} \end{bmatrix} \quad \text{for } j \rightarrow k \quad (8a)$$

$$\mathbf{M}^{\mathbf{v}, \mathbf{X}, \rho}(j) := \begin{bmatrix} \mathbf{v}_j & \mathbf{X}_j \\ \mathbf{X}_j^H & \rho_j \end{bmatrix} \quad \text{for } j \in \mathcal{V}. \quad (8b)$$

The branch flow model is defined in terms of variables $(\mathbf{s}, \mathbf{s}_{\Delta}, \mathbf{S}, \mathbf{v}, \ell, \mathbf{X}, \rho)$, and it is expressed as

$$\mathbf{v}_k = \mathbf{v}_j - (\mathbf{S}_{jk} z_{jk}^H + z_{jk} \mathbf{S}_{jk}^H) + z_{jk} \ell_{jk} z_{jk}^H \quad (9a)$$

$$\sum_{k:j \rightarrow k} \text{diag}(\mathbf{S}_{jk}) - \sum_{l:l \rightarrow j} \text{diag}(\mathbf{S}_{lj} - z_{lj} \ell_{lj}) \quad (9b)$$

$$= -\text{diag}(\mathbf{v}_j \mathbf{y}_j^H + \mathbf{X}_j \Gamma) + \mathbf{s}_j \quad (9c)$$

$$\mathbf{s}_{\Delta,j} = \text{diag}(\Gamma \mathbf{X}_j) \quad (9d)$$

$$\mathbf{v}_0 = \mathbf{V}_{\text{ref}} \mathbf{V}_{\text{ref}}^H \quad (9e)$$

$$\mathbf{M}^{\mathbf{v}, \mathbf{S}, \ell}(j, k) \succeq 0 \quad (9f)$$

$$\text{rank}(\mathbf{M}^{\mathbf{v}, \mathbf{S}, \ell}(j, k)) = 1 \quad (9g)$$

$$\mathbf{M}^{\mathbf{v}, \mathbf{X}, \rho}(j) \succeq 0 \quad (9h)$$

$$\text{rank}(\mathbf{M}^{\mathbf{v}, \mathbf{X}, \rho}(j)) = 1. \quad (9h)$$

Similar to BIM, (9g), (9h) are also derived from (2).

The AC-OPF problem in the BFM form can be formulated as:

$$\begin{aligned} & \text{minimize}_{\mathbf{s}, \mathbf{s}_{\Delta}, \mathbf{S}, \mathbf{v}, \ell, \mathbf{X}, \rho} f(\mathbf{s}, \mathbf{s}_{\Delta}) \end{aligned} \quad (10a)$$

$$\text{subject to} \quad (9), (3d) \quad (10b)$$

$$\text{diag}(\mathbf{V}_j \mathbf{V}_j^H) \leq \text{diag}(\mathbf{v}_j) \leq \text{diag}(\overline{\mathbf{V}_j \mathbf{V}_j^H}) \quad (10c)$$

III. ANALYTICAL RESULTS

The main challenge to solving OPF problems (6) and (10) is the nonconvex rank constraints in (5e), (5g), (9f) and (9h). If we drop all the rank-1 constraints, then we obtain

$$\begin{aligned} & \text{minimize}_{\mathbf{s}, \mathbf{s}_{\Delta}, \mathbf{W}, \mathbf{X}, \rho} f(\mathbf{s}, \mathbf{s}_{\Delta}) \end{aligned} \quad (11a)$$

$$\text{subject to} \quad (5a) - (5d), (5f), (3d), (6c). \quad (11b)$$

as the relaxation for the BIM and

$$\begin{aligned} & \text{minimize}_{\mathbf{s}, \mathbf{s}_{\Delta}, \mathbf{S}, \mathbf{v}, \ell, \mathbf{X}, \rho} f(\mathbf{s}, \mathbf{s}_{\Delta}) \end{aligned} \quad (12a)$$

$$\text{subject to} \quad (3d), (10c), (9a) - (9e), (9g) \quad (12b)$$

as the relaxation for the BFM. Solving the relaxed problems (11) and (12) could lead to solutions that are infeasible for the original nonconvex problems (6) and (10) respectively when the solutions do not satisfy the rank-1 constraints. In what follows, we will explore conditions under which optimal solutions of (6) and (10) can be recovered from their respective relaxations. First, the following lemma is presented, which is the main ingredient for subsequent results.

Lemma 1: Consider a block Hermitian matrix

$$\mathbf{M} := \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^H & \mathbf{C} \end{bmatrix} \quad (13)$$

where \mathbf{A} and \mathbf{C} are both square matrices. If $\mathbf{M} \succeq 0$ and $\mathbf{A} = \mathbf{x} \mathbf{x}^H$ for some vector \mathbf{x} , then there must exist some vector \mathbf{y} such that $\mathbf{B} = \mathbf{x} \mathbf{y}^H$.

Proof: As $\mathbf{M} \succeq 0$, it can be decomposed as

$$\mathbf{M} = \begin{bmatrix} \mathbf{M}_1 \\ \mathbf{M}_2 \end{bmatrix} \begin{bmatrix} \mathbf{M}_1^H & \mathbf{M}_2^H \end{bmatrix} \quad (14)$$

and $\mathbf{A} = \mathbf{M}_1 \mathbf{M}_1^H$, $\mathbf{B} = \mathbf{M}_1 \mathbf{M}_2^H$, $\mathbf{C} = \mathbf{M}_2 \mathbf{M}_2^H$. Because $\mathbf{A} = \mathbf{x} \mathbf{x}^H$ has rank-1, matrix \mathbf{M}_1 is in the column space of \mathbf{x} and has rank-1 as well. There must exist vector \mathbf{z} such that $\mathbf{M}_1 = \mathbf{x} \mathbf{z}^H$. As a result, $\mathbf{B} = \mathbf{M}_1 \mathbf{M}_2^H = \mathbf{x} \mathbf{z}^H \mathbf{M}_2^H = \mathbf{x} (\mathbf{M}_2 \mathbf{z})^H$. ■

One observation in Lemma 1 is when submatrices \mathbf{A} and \mathbf{B} are fixed and specified as $\mathbf{x} \mathbf{x}^H$ and $\mathbf{x} \mathbf{y}^H$, there are non-unique \mathbf{C} to make \mathbf{M} positive semi-definite. Similarly, in the relaxations (11) and (12), the optimal solutions are always non-unique. Taking (11) as an example, for any optimal solution $(\mathbf{s}^*, \mathbf{s}_{\Delta}^*, \mathbf{W}^*, \mathbf{X}^*, \rho^*)$, one could add to ρ^* an arbitrary positive semi-definite matrix to obtain a different optimal solution $(\mathbf{s}^*, \mathbf{s}_{\Delta}^*, \mathbf{W}^*, \mathbf{X}^*, \rho^* + \mathbf{K} \mathbf{K}^H)$. This non-uniqueness in the optimal ρ explains why in existing literature such as [22], the relaxation (11) could compute rank-1 \mathbf{W} (within numerical tolerance) but the resulting $\mathbf{M}^{\mathbf{W}, \mathbf{X}, \rho}$ is always not rank-1.

In fact, the next result shows in theory, if the optimal \mathbf{W} is perfectly of rank 1 without any numerical error, then an feasible and optimal solution of (6) is recoverable.

Theorem 1: If $\mathbf{u}^* = (\mathbf{s}^*, \mathbf{s}_\Delta^*, \mathbf{W}^*, \mathbf{X}^*, \rho^*)$ is an optimal solution to (11) that satisfies $\text{rank}(\mathbf{W}^*) = 1$, then a feasible and optimal solution of (6) can be recovered from \mathbf{u}^* .

Proof: We decompose \mathbf{W}_{jj}^* as $\mathbf{V}_j \mathbf{V}_j^H$ for each j , where \mathbf{V}_j is a vector. By Lemma 1, there exists vector $\mathbf{I}_{\Delta,j}$ such that $\mathbf{X}_j^* = \mathbf{V}_j \mathbf{I}_{\Delta,j}^H$. One could construct $\tilde{\rho}$ such that $\tilde{\rho}_j = \mathbf{I}_{\Delta,j} \mathbf{I}_{\Delta,j}^H$.

Since (11) is a relaxation of (6), for $(\mathbf{s}^*, \mathbf{s}_\Delta^*, \mathbf{W}^*, \mathbf{X}^*, \tilde{\rho})$ to be optimal for (6), it is sufficient that it is feasible for (6). Clearly, constraints (3d), (6c), (5a)–(5d) are satisfied because they are also the constraints in (11) and they do not involve the decision variable ρ . Constraint (5e) also holds as $\text{rank}(\mathbf{W}^*) = 1$. Further, by Lemma 1, we have

$$\begin{bmatrix} \mathbf{W}_{jj}^* & \mathbf{X}_j^* \\ (\mathbf{X}_j^*)^H & \tilde{\rho}_j \end{bmatrix} = \begin{bmatrix} \mathbf{V}_j \\ \mathbf{I}_{\Delta,j} \end{bmatrix} \begin{bmatrix} \mathbf{V}_j \\ \mathbf{I}_{\Delta,j} \end{bmatrix}^H$$

is both positive semi-definite and of rank-1. Hence, (5f) and (5g) are also satisfied. Hence, $(\mathbf{s}^*, \mathbf{s}_\Delta^*, \mathbf{W}^*, \mathbf{X}^*, \tilde{\rho})$ is feasible for (6), and this completes the proof. ■

Theorem 2: If $\mathbf{u}^* = (\mathbf{s}^*, \mathbf{s}_\Delta^*, \mathbf{S}^*, \mathbf{v}^*, \ell^*, \mathbf{X}^*, \rho^*)$ is an optimal solution to (12) and satisfies $\text{rank}(\mathbf{M}^{\mathbf{v}^*, \mathbf{S}^*, \ell^*}(j, k)) = 1$ for $j \sim k$ and $\text{rank}(\mathbf{v}_j^*) = 1$ for $j \in \mathcal{V}$, then an optimal solution of (10) can be recovered from \mathbf{u}^* .

The proof of Theorem 2 is omitted because it is similar to the proof of Theorem 1.

Theorem 1 asserts that in theory, the only critical non-convex constraint of (6) is (5e), in the sense that a solution satisfying (5g) could always be recovered whenever (5e) holds. However in practice, \mathbf{W}^* is typically not exactly rank-1 due to numerical precision and therefore Theorem 2 could not be directly applied to recover the optimal solution as long as numerical error exists. This is because the recovery method in Theorem 1 relies on the rank-1 decomposition of \mathbf{X}^* . In practice even if \mathbf{W}^* is close to being rank-1, the optimal \mathbf{X}^* could still be very different from being rank-1, as we will explain below in Remark 1.

Remark 1 (Spectrum Error): The matrix \mathbf{A} in (14) being approximately rank-1 does not necessarily mean that \mathbf{B} is also approximately rank-1.² For example, consider the case where \mathbf{x} , \mathbf{e}_1 , and \mathbf{e}_2 are orthogonal vectors with norms 1, 10^{-4} , and 10^{-5} , respectively. Similarly, let \mathbf{y} , \mathbf{z}_1 , and \mathbf{z}_2 be orthogonal vectors with norms 1, 10^4 , and 10^5 , respectively. Then, construct the matrix \mathbf{M} as in (14) with $\mathbf{M}_1 = [\mathbf{x} \ \mathbf{e}_1 \ \mathbf{e}_2]$ and $\mathbf{M}_2 = [\mathbf{y} \ \mathbf{z}_1 \ \mathbf{z}_2]$. Clearly, \mathbf{M} has the upper left diagonal block that is approximately rank-1. On the other hand, the upper right block is of rank 3 with three singular values of 1. Consequently, even when \mathbf{W}^* is close to rank-1 within a certain numerical tolerance, \mathbf{X}^* could be far from being a rank-1 matrix, especially if ρ^* already contains a large redundant positive semi-definite matrix $\mathbf{K}\mathbf{K}^H$. Decomposing \mathbf{X}^* as the product of two vectors, as in the proof of Theorem 1, could result in a large numerical error.

²Here, being approximately rank-1 means that the second largest eigenvalue of the matrix is nonzero but smaller than the largest eigenvalue by several orders of magnitude.

To summarize, there are two factors that prevent the relaxation output from being exact. The first is the non-uniqueness in the relaxation solution, and the second is that such non-uniqueness further greatly amplify the numerical error in computation. This finding motivates two algorithms for practical implementation.

A. Relaxation with Post-Processing

Remark 1 shows recovering the vector $\mathbf{I}_{\Delta,j}$ from \mathbf{X}_j^* can lead to poor numerical performance. In the first algorithm, we instead recover $\mathbf{I}_{\Delta,j}$ as $(\text{diag}(\Gamma \mathbf{V}_j))^{-1} \mathbf{s}_{\Delta,j}^*$ from (1b), and then we reconstruct $\tilde{\mathbf{X}}_j$ as $\mathbf{V}_j \mathbf{I}_{\Delta,j}^H$. If there is no numerical error, \mathbf{X}^* and $\tilde{\mathbf{X}}$ should be equal; however, in the presence of spectrum error, they could be different, as discussed in Remark 1. The pseudo code is provided in Algorithm 1.

Algorithm 1 Relaxation Algorithm with Post-Processing.

Input: $y, \mathcal{S}, \mathcal{S}_\Delta$

Output: Optimal solution $(\mathbf{s}, \mathbf{s}_\Delta, \mathbf{W}, \mathbf{X}, \rho)$ to (6).

- 1: Solve (11) to obtain $(\mathbf{s}^*, \mathbf{s}_\Delta^*, \mathbf{W}^*, \mathbf{X}^*, \rho^*)$.
 - 2: **if** $(\text{rank}(\mathbf{W}^*) > 1)$ **then**
 - 3: Output ‘Failed!’
 - 4: Exit
 - 5: **else**
 - 6: Decompose $\mathbf{W}_{jj}^* = \mathbf{V}_j \mathbf{V}_j^H$
 - 7: $\mathbf{I}_{\Delta,j} \leftarrow (\text{diag}(\Gamma \mathbf{V}_j))^{-1} \mathbf{s}_{\Delta,j}^*$
 - 8: $\tilde{\mathbf{X}} \leftarrow \mathbf{V}_j \mathbf{I}_{\Delta,j}^H, \tilde{\rho}_j \leftarrow \mathbf{I}_{\Delta,j} \mathbf{I}_{\Delta,j}^H$
 - 9: **return** $(\mathbf{s}^*, \mathbf{s}_\Delta^*, \mathbf{W}^*, \tilde{\mathbf{X}}, \tilde{\rho})$
 - 10: **end if**
-

Theorem 3: If Algorithm 1 does not fail, then its output is an optimal solution of (6).

Theorem 3 is the direct consequence of Theorem 1. Similarly for BFM, one could also apply post-processing to recover the solution of (10) from an optimal solution of (12). In the BFM, instead of checking the rank of \mathbf{W}^* , we check the rank of $\mathbf{M}^{\mathbf{v}^*, \mathbf{S}^*, \ell^*}(j, k)$ for each $j \sim k$ and \mathbf{v}_j^* for $j \in \mathcal{V}$.

B. Relaxation with Penalized Cost Function

The second algorithm we propose is to penalize and suppress the trace of ρ_j in the cost function. With such penalty term, the value of ρ^* will be unique for fixed \mathbf{W}^* and \mathbf{X}^* in the solution of (11) and the spectrum error can also be restricted. Similar penalization approaches were also previously proposed in [23], [24] to promote low-rank solutions. The penalized relaxed formulation under the BIM becomes

$$\begin{aligned} &\underset{\mathbf{s}, \mathbf{s}_\Delta, \mathbf{W}, \mathbf{X}, \rho}{\text{minimize}} && f(\mathbf{s}, \mathbf{s}_\Delta) + \lambda \sum_{j \in \mathcal{V}} \text{tr}(\rho_j) \end{aligned} \quad (15a)$$

$$\text{subject to} \quad (5a) - (5d), (5f), (3d), (6c) \quad (15b)$$

Similarly, the penalized relaxed program under the BFM becomes

$$\begin{aligned} &\underset{\mathbf{s}, \mathbf{s}_\Delta, \mathbf{S}, \mathbf{v}, \ell, \mathbf{X}, \rho}{\text{minimize}} && f(\mathbf{s}, \mathbf{s}_\Delta) + \lambda \sum_{j \in \mathcal{V}} \text{tr}(\rho_j) \end{aligned} \quad (16a)$$

$$\text{subject to} \quad (9a) - (9e), (9g), (3d), (10c). \quad (16b)$$

Because $\text{tr}(\rho_j)$ is linear and all constraints in (15b) and (16b) are convex, both (15) and (16) are convex optimization problems and can be efficiently solved in polynomial time. Here, $\lambda > 0$ controls the weight of $\sum \text{tr}(\rho_j)$ in the cost function. The pseudo code (based on BIM) is summarized in Algorithm 2. The algorithm for BFM is similar.

Algorithm 2 Relaxation Algorithm with Penalized Cost Function.

Input: y, S, S_Δ

Output: Optimal solution $(s, s_\Delta, W, X, \rho)$ to (6).

- 1: Pick a sufficiently small $\lambda > 0$
- 2: Solve (15) and obtain $u^* := (s^*, s_\Delta^*, W^*, X^*, \rho^*)$
- 3: **if** $(\text{rank}(W^*) > 1)$ **then**
- 4: Output ‘Failed!’
- 5: Exit
- 6: **else**
- 7: **return** u^*
- 8: **end if**

Because the cost function in the penalized program has been changed, the output of Algorithm 2 might not be the global optimal solution of (6). We next show that the output of Algorithm 2 serves as an approximation of the true optimal solution. We make the following assumption.

Assumption 1: The problem (11) has at least one finite optimal solution.

Now consider a sequence of positive and decreasing λ_i for $i = 1, 2, \dots$ such that $\lambda_i \rightarrow 0$ as $i \rightarrow \infty$. Taking BIM as an example, let the optimal solution of (15) with respect to λ_i be $u^{(i)}$.³ Then the following lemma implies the sequence $u^{(i)}$ has a limit point.

Lemma 2: The sequence $(u^{(i)})_{i=1}^\infty$ resides in a compact set, and hence has a limit point.

Proof: Because all the constraints in (15) are closed, we only need to prove boundedness. By assumption, $s_j^{(i)}$ and $s_{\Delta,j}^{(i)}$ at bus j are in compact sets S_j and $S_{\Delta,j}$ respectively. The positive semi-definite matrix $W^{(i)}$ has upper bounds on its diagonal elements and is therefore bounded. We only need to show that $\sum_j \text{tr}(\rho_j^{(i)})$ is also bounded because the boundedness of $X^{(i)}$ is implied by the constraint (5f) as long as $\sum_j \text{tr}(\rho_j^{(i)})$ is bounded.

To show $\sum_j \text{tr}(\rho_j^{(i)})$ is bounded, let $\hat{u} = (\hat{s}, \hat{s}_\Delta, \hat{W}, \hat{X}, \hat{\rho})$ be an optimal solution of (11). Then \hat{u} is feasible for (15) regardless of the value of λ . For any i , we must have $\sum_j \text{tr}(\rho_j^{(i)}) \leq \sum_j \text{tr}(\hat{\rho}_j)$; otherwise, \hat{u} will always give a strictly smaller cost value in (15) for $\lambda = \lambda_i$ and it would contradict the optimality of $u^{(i)}$. ■

Suppose $\tilde{u} := (\tilde{s}, \tilde{s}_\Delta, \tilde{W}, \tilde{X}, \tilde{\rho})$ is an arbitrary limit point of the sequence $u^{(i)}$. We present sufficient conditions for \tilde{u} to be an optimal solution of (6).

Theorem 4: If $\text{rank}(\tilde{W}) = 1$, then \tilde{u} is a globally optimal solution of (6).

The proof of Theorem 4 can be found in [25]. Theorem 4 shows that when we solve the penalized program with a

sequence of decreasing λ_i that converge to 0, any limit point would be a global optimal for (6) as long as the W matrix associated with the limit point is of rank-1. In our simulations, we apply Algorithm 2 to solve (15) with a fixed but sufficiently small λ , which usually results in rank-1 solutions.

Remark 2: Further, if all optimal solutions of (11) have the same value for s, s_Δ, W, X , then Algorithm 1 succeeds if and only if $\text{rank}(\tilde{W}) = 1$ holds in Theorem 4. If Algorithm 1 succeeds, its output will also be the same as \tilde{u} .

IV. MODEL EQUIVALENCE

In the previous sections, our results for the BIM and BFM always come in pairs and are analogous. In single-phase networks and multi-phase systems with only wye connections, [26] and [19] have shown that the two models are equivalent in the sense that one will produce an exact solution if and only if the other will. We show in this subsection that a similar result holds in the presence of delta connections.

Definition 1: Consider two optimization problems

$$\underset{x}{\text{minimize}} \quad f_A(x) \quad \text{subject to} \quad x \in \mathcal{X} \quad (17)$$

$$\underset{y}{\text{minimize}} \quad f_B(y) \quad \text{subject to} \quad y \in \mathcal{Y}. \quad (18)$$

We say (17) and (18) are *equivalent* if there exist mappings $g_1 : \mathcal{X} \rightarrow \mathcal{Y}$ and $g_2 : \mathcal{Y} \rightarrow \mathcal{X}$ such that

$$\begin{aligned} x \in \mathcal{X} &\Rightarrow g_1(x) \in \mathcal{Y}, \quad f_A(x) = f_B(g_1(x)), \\ y \in \mathcal{Y} &\Rightarrow g_2(y) \in \mathcal{X}, \quad f_B(y) = f_A(g_2(y)). \end{aligned}$$

We do not require g_1 and g_2 to be bijections, but if one of the mappings is a bijection, then we can always set the other as its inverse. We denote the decision variables for the BIM as

$$u^{\text{BIM}} = (s^{\text{BIM}}, s_\Delta^{\text{BIM}}, W^{\text{BIM}}, X^{\text{BIM}}, \rho^{\text{BIM}})$$

and the decision variables for the BFM as

$$u^{\text{BFM}} = (s^{\text{BFM}}, s_\Delta^{\text{BFM}}, S^{\text{BFM}}, v^{\text{BFM}}, \ell^{\text{BFM}}, X^{\text{BFM}}, \rho^{\text{BFM}}).$$

The superscripts here are to distinguish the same variable for different models.

Proposition 1: Problems (11) and (12) are equivalent. Moreover, for the pairs g_1 and g_2 in Definition 1, if u^{BIM} satisfies (5e), then $g_1(u^{\text{BIM}})$ satisfies (9f). If u^{BFM} satisfies (9f), then $g_2(u^{\text{BFM}})$ satisfies (5e).

Note that (11) and (12) are the relaxed BIM and BFM models. The proposition above implies that besides the equivalence between nonconvex BIM and BFM models as we derived in Section II, their relaxations are also equivalent. We only sketch a proof here by providing the mappings g_1 and g_2 , where g_1 can be written as

$$s^{\text{BFM}} = s^{\text{BIM}}, \quad s_\Delta^{\text{BFM}} = s_\Delta^{\text{BIM}} \quad (19a)$$

$$S_{jk}^{\text{BFM}} = (W_{jj}^{\text{BIM}} - W_{jk}^{\text{BIM}})y_{jk}^H \quad (19b)$$

$$v_j^{\text{BFM}} = W_{jj}^{\text{BIM}} \quad (19c)$$

$$\ell_{jk}^{\text{BFM}} = y_{jk}(W_{jj}^{\text{BIM}} + W_{kk}^{\text{BIM}} - W_{jk}^{\text{BIM}} - W_{kj}^{\text{BIM}})y_{jk}^H \quad (19d)$$

$$X_j^{\text{BFM}} = X_j^{\text{BIM}}, \quad \rho_j^{\text{BFM}} = \rho_j^{\text{BIM}} \quad (19e)$$

³If the program has multiple solutions, then pick any one of them.

and g_2 as

$$\mathbf{s}^{\text{BIM}} = \mathbf{s}^{\text{BFM}}, \mathbf{s}_{\Delta}^{\text{BIM}} = \mathbf{s}_{\Delta}^{\text{BFM}} \quad (20a)$$

$$\mathbf{W}_{jj}^{\text{BIM}} = \mathbf{v}_j^{\text{BFM}} \quad (20b)$$

$$\mathbf{W}_{jk}^{\text{BIM}} = \begin{cases} \mathbf{v}_j^{\text{BFM}} - \mathbf{S}_{jk}^{\text{BFM}} \mathbf{z}_{jk}^{\text{H}}, & \text{if } j \rightarrow k \\ (\mathbf{W}_{kj}^{\text{BIM}})^{\text{H}}, & \text{if } k \rightarrow j \end{cases} \quad (20c)$$

$$\mathbf{X}_j^{\text{BIM}} = \mathbf{X}_j^{\text{BFM}}, \rho_j^{\text{BIM}} = \rho_j^{\text{BFM}}. \quad (20d)$$

For g_2 , the value of $\mathbf{W}_{jk}^{\text{BIM}}$ where $j \neq k$ and $(j, k) \notin \mathcal{E}$ can be determined arbitrarily as long as $\mathbf{W} \succeq 0$. As \mathcal{G} is a tree, we can always complete the matrix $\mathbf{W}^{\text{BIM}} \succeq 0$, but not necessarily in a unique way.

Proposition 1 shows that to apply Algorithm 1, if an optimal solution of (11) can produce an exact solution of (6), then there must also be an optimal solution of (12) that can produce an exact solution of (10), even though both (6) and (10) may have multiple solutions. The converse is also true. Informally, for Algorithm 1, both the BIM and BFM have the same capability of producing exact solutions.

The same holds for the penalized program. The next proposition can be easily proved using the same mappings g_1 and g_2 in (19) and (20), respectively.

Proposition 2: Problems (15) in the BIM and (16) in the BFM are equivalent when λ takes the same value for both problems.

Though BIM and BFM are mathematically equivalent, the two models may behave differently in practice and shed lights on different properties. Some analysis may rely on the structure of one model but not the other, which is indeed the case for single-phase networks [27]. The equivalence implies that one could freely choose a model that is more convenient for a specific problem.

V. NUMERICAL RESULTS

In this section, we show the ability of the proposed relaxation algorithms to recover the optimal solution to (6) and (10). We use the IEEE 13-, 37-, and 123-node distribution feeders [28] to assess the exactness of both algorithms for both the BIM and BFM models. Note that the IEEE 123-bus feeder does not include delta-connected components. Hence, we artificially added 4 delta-connected loads to the feeder to assess the performance of the proposed approaches. Therefore, all feeders in our simulation will include delta connections. In our experiments, we check how close the output matrices $\mathbf{W}, \mathbf{M}^{\mathbf{W}, \mathbf{X}, \rho}, \mathbf{M}^{\mathbf{V}, \mathbf{S}, \ell}, \mathbf{M}^{\mathbf{V}, \mathbf{X}, \rho}$ are to being rank-1, and we evaluate the maximum violation of the constraints when the decision variables are produced from the two proposed algorithms. For all the experiments in this section, we show that both algorithms succeed up to numerical precision, and each has its own advantages and disadvantages.

In previous sections, when we refer to Algorithm 1 and 2 as being exact, the claim is in the sense that Algorithm 1 in theory would produce the globally optimal solution of (6) if there were no numerical error (i.e., Theorem 3), and the output of Algorithm 2 would converge to the globally optimal solution of (6) as λ goes to 0 (i.e., Theorem 4). In practice, the machine always has finite precision and we always set λ

as a fixed small number in the program. Therefore, the output cost of Algorithm 1 in our simulation should be regarded as a lower bound of the globally optimal cost and the cost of Algorithm 2 should be regarded as an approximation. With higher precision and smaller λ (depending on the precision), the output cost of both algorithms will be closer to the actual globally optimal cost.

A. Experimental Setup

The load transformer in the IEEE test feeders are modeled as lines with equivalent impedance, whereas the substation transformers and regulators are removed. The switches are assumed to be open or short according to their default status. The capacitor banks are modeled as controllable reactive power sources with continuous control space. The same modification is also commonly applied in the literature [18], [22].

The voltage at the substation is assumed to be $\mathbf{V}_{\text{ref}} = \bar{V}[1, e^{-j2\pi/3}, e^{j2\pi/3}]^T$, where \bar{V} is the maximum allowed voltage magnitude. The operational constraints for controllable loads are set as in [22]. The AC-OPF problem is solved with the cost function $f(\mathbf{s}, \mathbf{s}_{\Delta})$ comprising three parts.⁴ The first part minimizes the total power losses in the network, and it can be written as

$$p_{\text{loss}} = \sum_{j \in \mathcal{V}} \sum_{\phi \in \Phi_Y^j} \mathbf{p}_j^{\phi} + \sum_{j \in \mathcal{V}} \sum_{\phi \in \Phi_{\Delta}^j} \mathbf{p}_{\Delta,j}^{\phi}.$$

The second part penalizes deviations of the active and reactive injection profile from nominal profiles, and it is given by $d_p(\mathbf{p}, \mathbf{p}_{\Delta})$ and $d_q(\mathbf{q}, \mathbf{q}_{\Delta})$ as follows.

$$d_p(\mathbf{p}, \mathbf{p}_{\Delta}) = \sum_{j \in \mathcal{V}} \sum_{\phi \in \Phi_Y^j} \frac{(\mathbf{p}_j^{\phi} - \bar{\mathbf{p}}_j^{\phi})^2}{2\bar{\mathbf{p}}_j^{\phi}} + \sum_{j \in \mathcal{V}} \sum_{\phi \in \Phi_{\Delta}^j} \frac{(\mathbf{p}_{\Delta,j}^{\phi} - \bar{\mathbf{p}}_{\Delta,j}^{\phi})^2}{2\bar{\mathbf{p}}_{\Delta,j}^{\phi}},$$

$$d_q(\mathbf{q}, \mathbf{q}_{\Delta}) = \sum_{j \in \mathcal{V}} \sum_{\phi \in \Phi_Y^j} \frac{(\mathbf{q}_j^{\phi} - \bar{\mathbf{q}}_j^{\phi})^2}{2\bar{\mathbf{q}}_j^{\phi}} + \sum_{j \in \mathcal{V}} \sum_{\phi \in \Phi_{\Delta}^j} \frac{(\mathbf{q}_{\Delta,j}^{\phi} - \bar{\mathbf{q}}_{\Delta,j}^{\phi})^2}{2\bar{\mathbf{q}}_{\Delta,j}^{\phi}}.$$

The values $\bar{\mathbf{p}}_j^{\phi}, \bar{\mathbf{q}}_j^{\phi}, \bar{\mathbf{p}}_{\Delta,j}^{\phi}, \bar{\mathbf{q}}_{\Delta,j}^{\phi}$ represent the nominal active and reactive injection values for phase ϕ at bus j . All the tracking errors are normalized by their nominal values to have the same order of magnitude for all quantities. In addition, $\Phi_Y^j \subseteq \{a, b, c\}$ and $\Phi_{\Delta}^j \subseteq \{ab, bc, ca\}$ denote the available wye and delta connections at bus $j \in \mathcal{V}$, respectively. Penalizing the deviation of power injection can characterize either the operational cost of controllable loads, the curtailment of photovoltaic systems, or the charging cost of batteries. The same cost expression was also used in [22], [29].

The last part minimizes the deviation of the power injections at the substation from the reference injections $\bar{p}_0, \bar{q}_0 \in \mathbb{R}$ provided by the transmission system operator. Therefore, the

⁴We will use \mathbf{p}, \mathbf{q} to denote the real and imaginary parts of \mathbf{s} , and $\mathbf{p}_{\Delta}, \mathbf{q}_{\Delta}$ to denote the real and imaginary parts of \mathbf{s}_{Δ} .

system operational cost function can be written as

$$f(\mathbf{s}, \mathbf{s}_\Delta) = \mu_\ell p_{\text{loss}} + w_p d_p(\mathbf{p}, \mathbf{p}_\Delta) + w_q d_q(\mathbf{q}, \mathbf{q}_\Delta) + \mu_p \frac{(\mathbf{1}^\top \mathbf{p}_0 - \bar{p}_0)^2}{\bar{p}_0} + \mu_q \frac{(\mathbf{1}^\top \mathbf{q}_0 - \bar{q}_0)^2}{\bar{q}_0}.$$

The nonnegative weights w_p , w_q , μ_ℓ , μ_p and μ_q are used to reflect the relative importance of the components of the cost function and are set as follows

$$w_p = w_q = \mu_\ell = 1, \quad \mu_p = \mu_q = 4. \quad (21)$$

B. Exactness Results for Algorithm 1

In this subsection, we assess the quality of the solutions recovered using Algorithm 1. We solve (11) for the BIM as well as (12) for the BFM with different values of voltage limits for the three considered feeders. We invoke the Mosek 8.0 conic solver using CVX, a MATLAB-based convex optimization toolbox.

The left-hand side of Table I provides the result of Algorithm 1 based on the BIM. The voltage column represents the maximum and minimum voltage deviation allowed, i.e., 3% means that the value of \bar{V} and \underline{V} are set to 1.03 pu and 0.97 pu, respectively. We assess the rank of matrices \mathbf{W}_{jj} , for all $j \in \mathcal{V}$, in terms of the ratio between the top two largest eigenvalues of these matrices. The maximum ratio among all $j \in \mathcal{V}$ is listed in the table. In the solution of (11) (before post-processing), the ratio between the two maximum eigenvalues of the matrices $\mathbf{M}^{\mathbf{W}, \mathbf{X}, \rho}$ is on the order of 10^{-1} , and after the post-processing in Algorithm 1, the final $\mathbf{M}^{\mathbf{W}, \mathbf{X}, \rho}$ -ratio will be dominated by \mathbf{W} -ratio and is thus not informative to be displayed in the table. Because of the spectrum error, the output $\tilde{\mathbf{X}}$ could be different from \mathbf{X}^* , and thus having a very small \mathbf{W} -ratio is not enough to guarantee the feasibility of the final output of Algorithm 1. Therefore, we also assess the infeasibility of the power flow equations by measuring the maximum violation in (5a) for the solutions returned by Algorithm 1. Here, the violation is defined as the difference between the left- and right-hand sides of (5a) (in kW) when $\mathbf{s}, \mathbf{W}, \mathbf{X}$ are evaluated as the output of Algorithm 1. In our simulations, the infeasibility is on the order of 10^{-3} to 10^{-1} kW, which reflects the effect of the spectrum error after the post-processing. As a benchmark, the load injections for those feeders are on the order of 10^1 to 10^2 kW, and are typically two orders of magnitude higher than the infeasibility.

On the right-hand side of Table I, the rank of $\mathbf{M}^{\mathbf{V}, \mathbf{S}, \ell}(j, k)$ for all lines $(j, k) \in \mathcal{E}$ is examined for the same algorithm under the BFM. Again, we present the maximum ratio between the two largest eigenvalues. Similar to the BIM, the infeasibility, i.e., the violation of (1a), is shown in the table.

C. Exactness Results for Algorithm 2

In our setting, the penalized formulations (15) and (16) are solved with the parameter $\lambda = 10$ in all experiments. We will later show how the value of λ affects the solution quality.

Table II presents the maximum ratio between the top two largest eigenvalues of $\mathbf{M}^{\mathbf{W}, \mathbf{X}, \rho}(j)$ for the BIM and $\mathbf{M}^{\mathbf{V}, \mathbf{S}, \ell}(j)$

for the BFM returned by the solvers. Comparing the infeasibility of the solutions obtained using Algorithm 1, shown in Table I, and Algorithm 2, shown in Table II, it is clear that adding a penalty helps reduce the effect of the spectrum error and leads to globally optimal solutions with much lower infeasibility.

To assess the effect of the penalization approach on the quality of the solutions in terms of cost and feasibility, Table III shows the effect of increasing the penalty parameter in the cost function as well as the maximum infeasibility of the power equations (in kVA) for the IEEE 37-bus network with 3% voltage limits. The cost in Table III is evaluated without the penalty term. Note that the case $\lambda = 0$ corresponds to the output of Algorithm 1. Although the solution feasibility is enhanced by increasing the penalty parameter, the cost associated with the solution obtained also increases. Note that the cost obtained with $\lambda = 0$, i.e., from Algorithm 1, represents a lower bound for the optimal cost of the original AC-OPF problem.⁵ In addition, increasing the penalty parameter beyond the values considered in Table III leads to uninteresting solutions because the cost function becomes dominated by the penalty term. For real applications, we suggest to use binary search to find the smallest λ such that the infeasibility of the solution is within the user-specified tolerance range.

D. Results with Distributed Energy Resources

We now assess the performance of the proposed approach in a more general setting where distributed energy resources (DERs), such as photovoltaics (PV), are installed. In this simulation, we utilize the IEEE 37-bus distribution feeder where we assume that five PV systems are installed in delta-connections at the buses 725, 729, 731, 732, and 740. The available power at these units is set at 120, 75, 90, 105, and 180 KW, respectively. We also assume that all the PV inverters can provide reactive power support such that the resultant power factor is at least 0.8. Using this modified feeder, we evaluate the performance of the proposed algorithms when the cost function is p_{loss} , i.e., $\mu_p = \mu_q = w_p = w_q = 0$ and $\mu_\ell = 1$. In addition, we set the upper and lower bounds on voltage magnitudes to be 1.03 pu and 0.97 pu, respectively, in this simulation. In Table IV, the results of both Algorithm 1 and Algorithm 2 are presented. It is consistent with previous sections that Algorithm 2 often has lower infeasibility compared to Algorithm 1. To assess the voltage magnitudes resulting from the proposed algorithms, Fig. 1 depicts the voltage magnitude at all phases for the solution produced by Algorithm 2 ($\lambda = 1$). It is worth noting that the same voltage profile is obtained by both the BFM and BIM formulations. The results confirms that the voltages in the solution are within the operational limits.

In addition, we evaluate the performance of both the BIM and BFM under a different cost function which also includes the substation power deviation, i.e., $w_p = w_q = 0$, $\mu_p =$

⁵Here is the reason why the cost of Algorithm 1 is regarded as a lower bound. One consequence of having the spectrum error is that numerical error in BIM formulation could lead to larger constraint violation (as indicated by the infeasibility). Therefore, the output cost (with slight constraint violation) may be lower than the actual optimal cost within the feasible set. We would expect the actual optimal cost to be exact if there were no numerical error.

TABLE I
RANK AND INFEASIBILITY FOR THE OUTPUTS OF ALGORITHM 1 (WITH POST-PROCESSING).

Network	# of Δ -loads	Voltage	BIM		BFM	
			W-ratio	Infeas. (kW)	$M^{v,S,\ell}$ -ratio	Infeas. (kW)
IEEE-13	2	3%	1.91×10^{-8}	3.20×10^{-1}	1.74×10^{-5}	4.29×10^{-2}
		5%	2.73×10^{-9}	3.20×10^{-1}	1.58×10^{-5}	4.10×10^{-2}
IEEE-37	25	3%	4.81×10^{-10}	9.84×10^{-2}	7.81×10^{-5}	9.67×10^{-2}
		5%	2.77×10^{-8}	9.83×10^{-2}	2.70×10^{-5}	9.72×10^{-2}
IEEE-123	4	3%	7.75×10^{-8}	1.54×10^{-3}	1.07×10^{-4}	1.03×10^{-2}
		5%	7.67×10^{-8}	1.54×10^{-3}	9.64×10^{-5}	1.02×10^{-2}

TABLE II
RANK AND INFEASIBILITY FOR THE OUTPUTS OF ALGORITHM 2 (WITH PENALIZED COST FUNCTION).

Network	Voltage	BIM			BFM		
		W-ratio	$M^{W,X,\rho}$ -ratio	Infeas. (kW)	$M^{v,S,\ell}$ -ratio	$M^{v,X,\rho}$ -ratio	Infeas. (kW)
IEEE-13	3%	1.34×10^{-10}	2.36×10^{-9}	8.85×10^{-2}	1.44×10^{-10}	1.97×10^{-10}	4.43×10^{-5}
	5%	1.31×10^{-10}	1.96×10^{-9}	8.84×10^{-2}	1.36×10^{-10}	1.57×10^{-10}	1.46×10^{-5}
IEEE-37	3%	3.04×10^{-8}	6.22×10^{-8}	5.75×10^{-6}	8.85×10^{-8}	3.38×10^{-5}	1.45×10^{-6}
	5%	2.94×10^{-8}	1.05×10^{-8}	1.06×10^{-6}	2.12×10^{-8}	3.18×10^{-5}	1.00×10^{-6}
IEEE-123	3%	1.45×10^{-9}	7.03×10^{-9}	9.13×10^{-7}	1.05×10^{-8}	8.99×10^{-9}	1.34×10^{-6}
	5%	1.94×10^{-8}	9.31×10^{-8}	7.84×10^{-6}	7.98×10^{-9}	6.59×10^{-9}	1.40×10^{-6}

TABLE III
EFFECT OF THE PENALTY PARAMETER ON THE COST AND INFEASIBILITY.

λ	BIM		BFM	
	Cost	Infeas. (kW)	Cost	Infeas. (kW)
0	100.0036	9.84×10^{-2}	100.0194	9.67×10^{-2}
0.1	103.9504	1.15×10^{-2}	104.7141	5.99×10^{-4}
1	104.7846	6.00×10^{-5}	104.7840	1.80×10^{-5}
10	104.7886	5.75×10^{-6}	104.7982	1.45×10^{-6}
100	104.9431	3.17×10^{-6}	105.0332	1.23×10^{-7}

TABLE IV
RESULTS ON IEEE 37-BUS NETWORK WITH DERS INSTALLED WHEN MINIMIZING THE ELECTRICAL LOSSES.

Model	Algorithm 1		Algorithm 2 ($\lambda = 1$)	
	Cost	Infeas. (kW)	Cost	Infeas. (kW)
BIM	2.5571	6.68×10^{-2}	4.0415	4.42×10^{-6}
BFM	2.5631	6.57×10^{-2}	4.0442	7.00×10^{-6}

$\mu_q = 4$, and $\mu_\ell = 1$. Furthermore, we set the reference substation injection such that it is achievable only if the available PV power is curtailed. Therefore, the reference power tracking term in the cost function becomes not increasing in

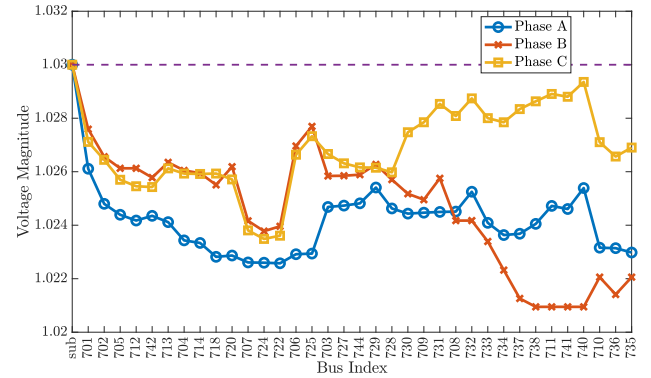


Fig. 1. The voltage magnitude at all buses for the solution obtained through Algorithm 2 ($\lambda = 1$) for the modified IEEE 37-bus feeder.

power injections. This is known to lead to inexactness of the relaxation. We test this cost function using Algorithm 1 and Algorithm 2 ($\lambda = 1$). Table V shows the infeasibility and the cost function of the solutions obtained using both BIM and BFM. We can see that the infeasibility of the solution obtained using Algorithm 1 is aggravated due to the use of the cost function that is not increasing with power injections. However, adding the penalty term in Algorithm 2 is enough to significantly reduce the infeasibility.

TABLE V
RESULTS ON IEEE 37-BUS NETWORK WITH DERS INSTALLED WHEN
MINIMIZING REFERENCE TRACKING COST AND ELECTRICAL LOSS.

Model	Algorithm 1		Algorithm 2 ($\lambda = 1$)	
	Cost	Infeas. (kW)	Cost	Infeas. (kW)
BIM	18.2124	1.42×10^{-1}	19.6029	6.00×10^{-7}
BFM	18.2281	1.40×10^{-1}	19.6026	6.35×10^{-6}

TABLE VI
COMPUTATIONAL TIME FOR BOTH BIM AND BFM (IN SECONDS).

Model	IEEE-13	IEEE-37	IEEE-123
BIM	2.46	5.56	9.03
BFM	2.77	4.93	9.87

E. Algorithm Summary and Comparison

Algorithms 1 and 2 can be useful for different applications. Algorithm 1 solves the un-penalized problem and therefore prioritizes cost minimization at the cost of larger constraint violation. The simulation shows that the infeasibility is typically two orders of magnitude smaller than the load injections and should be acceptable. Algorithm 2, on the other hand, can recover a solution with much smaller constraint violation, but the optimal cost is higher because of the penalty term.

The simulation results also show that the methods under BFM are more numerically stable than BIM, in terms of the infeasibility in Table I and II. This observation is consistent with the performance of two models for single-phase feeders and feeders without delta-connections, as shown in [19], [10].

We also benchmark the computational time of the proposed algorithms in our case studies. Since both algorithms require solving similar optimization problems with slightly different cost functions, the computational time of the two algorithms is similar. Hence, we only present the computational time of Algorithm 1 for all networks using both the BIM and BFM formulations. The algorithm was implemented using Mosek 8.0 as a conic solver on a laptop with Intel Core i9 CPU (2.40 GHz), 16 GB RAM, macOS Catalina OS, and MATLAB R2019b. The results show that the proposed algorithms take less than 10 seconds to solve the AC-OPF problem for the IEEE 123-bus network on a standard laptop, which demonstrates the computational efficiency of the proposed algorithms. More advanced methods such as sparse semi-definite programming solvers, e.g., [30], can further scale the implementation to thousands of buses.

VI. CONCLUSION

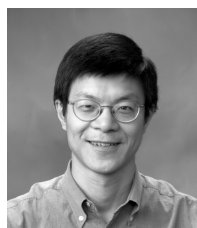
This paper studied the SDP relaxation of the AC-OPF problem for an unbalanced three-phase radial network with delta connections, formulated under both the BIM and BFM. We showed the equivalence between the BIM and BFM formulations and presented sufficient conditions for recovering

exact solutions of the nonconvex AC-OPF formulations from their respective relaxations. The paper also showed why conventional relaxation (by directly dropping rank-1 constraints) always fails when the sufficient conditions are approximately satisfied. It is due to the non-uniqueness in the relaxation solution and the spectrum error in computation. Inspired by this finding, we then proposed two algorithms which are guaranteed to produce exact solutions whenever our sufficient conditions are satisfied. One applies post-processing and produces lower cost but larger constraint violation. The other adds a penalty term and produces higher cost but smaller constraint violation. In simulations, we demonstrated that for three IEEE standard test cases, both algorithms are able to recover near globally optimal solutions with tolerable constraint violation and cost suboptimality.

REFERENCES

- [1] J. Carpentier, "Contribution to the economic dispatch problem," *Bulletin de la Societe Francoise des Electriciens*, vol. 3, no. 8, pp. 431–447, 1962.
- [2] D. Bienstock and A. Verma, "Strong NP-hardness of AC power flows feasibility," *Operations Research Letters*, vol. 47, no. 6, pp. 494–501, 2019.
- [3] K. Lehmann, A. Grastien, and P. Van Hentenryck, "AC-feasibility on tree networks is NP-hard," *IEEE Transactions on Power Systems*, vol. 31, no. 1, pp. 798–801, 2016.
- [4] J. A. Momoh, R. Adapa, and M. El-Hawary, "A review of selected optimal power flow literature to 1993. I. nonlinear and quadratic programming approaches," *IEEE Transactions on Power Systems*, vol. 14, no. 1, pp. 96–104, 1999.
- [5] D. Wells, "Method for economic secure loading of a power system," in *Proceedings of the Institution of Electrical Engineers*, vol. 115, no. 8. IET, 1968, pp. 1190–1194.
- [6] R. Burchett, H. Happ, and K. Wirgau, "Large scale optimal power flow," *IEEE Transactions on Power Apparatus and Systems*, no. 10, pp. 3722–3732, 1982.
- [7] A. Bernstein, C. Wang, E. Dall'Anese, J.-Y. Le Boudec, and C. Zhao, "Load flow in multiphase distribution networks: Existence, uniqueness, non-singularity and linear models," *IEEE Transactions on Power Systems*, vol. 33, no. 6, pp. 5832–5843, 2018.
- [8] R. A. Jabr, "Radial distribution load flow using conic programming," *IEEE Transactions on Power Systems*, vol. 21, no. 3, pp. 1458–1459, 2006.
- [9] X. Bai, H. Wei, K. Fujisawa, and Y. Wang, "Semidefinite programming for optimal power flow problems," *International Journal of Electrical Power & Energy Systems*, vol. 30, no. 6-7, pp. 383–392, 2008.
- [10] D. K. Molzahn and I. A. Hiskens, "A survey of relaxations and approximations of the power flow equations," *Foundations and Trends® in Electric Energy Systems*, vol. 4, no. 1-2, pp. 1–221, 2019.
- [11] S. H. Low, "Convex relaxation of optimal power flow—part II: Exactness," *IEEE Transactions on Control of Network Systems*, vol. 1, no. 2, pp. 177–189, 2014.
- [12] H. Hijazi, C. Coffrin, and P. Van Hentenryck, "Convex quadratic relaxations for mixed-integer nonlinear programs in power systems," *Mathematical Programming Computation*, vol. 9, no. 3, pp. 321–367, 2017.
- [13] A. S. Zamzam, N. D. Sidiropoulos, and E. Dall'Anese, "Beyond relaxation and Newton-Raphson: Solving AC OPF for multi-phase systems with renewables," *IEEE Transactions on Smart Grid*, vol. 9, no. 5, pp. 3966–3975, 2018.
- [14] A. S. Zamzam, C. Zhao, E. Dall'Anese, and N. D. Sidiropoulos, "A QCQP approach for OPF in multiphase radial networks with wye and delta connections," in *Proceedings of the 10th Bulk Power Systems Dynamics and Control Symposium*, Espinho, Portugal, 2017.
- [15] R. Berg, E. Hawkins, and W. Pleines, "Mechanized calculation of unbalanced load flow on radial distribution circuits," *IEEE Transactions on power apparatus and systems*, no. 4, pp. 415–421, 1967.
- [16] M. Laughton, "Analysis of unbalanced polyphase networks by the method of phase co-ordinates. part 1: System representation in phase frame of reference," in *Proceedings of the Institution of Electrical Engineers*, vol. 115, no. 8. IET, 1968, pp. 1163–1172.

- [17] R. Madani, S. Sojoudi, and J. Lavaei, "Convex relaxation for optimal power flow problem: Mesh networks," *IEEE Transactions on Power Systems*, vol. 30, no. 1, pp. 199–211, 2014.
- [18] E. Dall'Anese, H. Zhu, and G. B. Giannakis, "Distributed optimal power flow for smart microgrids," *IEEE Transactions on Smart Grid*, vol. 4, no. 3, pp. 1464–1475, 2013.
- [19] L. Gan and S. H. Low, "Convex relaxations and linear approximation for optimal power flow in multiphase radial networks," in *2014 Power Systems Computation Conference*. IEEE, 2014, pp. 1–9.
- [20] F. Zhou, Y. Chen, and S. H. Low, "Sufficient conditions for exact semidefinite relaxation of optimal power flow in unbalanced multiphase radial networks," in *2019 IEEE 58th Conference on Decision and Control (CDC)*. IEEE, 2019, pp. 6227–6233.
- [21] W. Wang and N. Yu, "Chordal conversion based convex iteration algorithm for three-phase optimal power flow problems," *IEEE Transactions on Power Systems*, vol. 33, no. 2, pp. 1603–1613, 2017.
- [22] C. Zhao, E. Dall'Anese, and S. H. Low, "Optimal power flow in multiphase radial networks with delta connections," National Renewable Energy Lab. (NREL), Golden, CO (United States), Tech. Rep., 2017.
- [23] R. Madani, M. Ashraphijuo, and J. Lavaei, "Promises of conic relaxation for contingency-constrained optimal power flow problem," *IEEE Transactions on Power Systems*, vol. 31, no. 2, pp. 1297–1307, 2015.
- [24] D. K. Molzahn, C. Jozs, I. A. Hiskens, and P. Panciatici, "A Laplacian-based approach for finding near globally optimal solutions to OPF problems," *IEEE Transactions on Power Systems*, vol. 32, no. 1, pp. 305–315, 2017.
- [25] F. Zhou, A. S. Zamzam, S. H. Low, and N. D. Sidiropoulos, "Exactness of opf relaxation on three-phase radial networks with delta connections," *arXiv preprint arXiv:2005.07803*, 2020.
- [26] S. Bose, S. H. Low, T. Teerarattul, and B. Hassibi, "Equivalent relaxations of optimal power flow," *IEEE Transactions on Automatic Control*, vol. 60, no. 3, pp. 729–742, 2014.
- [27] S. H. Low, "Convex relaxation of optimal power flow-part I: Formulations and equivalence," *IEEE Transactions on Control of Network Systems*, vol. 1, no. 1, pp. 15–27, 2014.
- [28] K. Schneider, B. Mather, B. Pal, C.-W. Ten, G. Shirek, H. Zhu, J. Fuller, J. Pereira, L. Ochoa, L. De Araujo *et al.*, "Analytic considerations and design basis for the IEEE distribution test feeders," *IEEE Transactions on power systems*, vol. 33, no. 3, pp. 3181–3188, 2017.
- [29] E. Dall'Anese, S. S. Guggilam, A. Simonetto, Y. C. Chen, and S. V. Dhople, "Optimal regulation of virtual power plants," *IEEE Transactions on Power Systems*, vol. 33, no. 2, pp. 1868–1881, 2017.
- [30] R. Y. Zhang and J. Lavaei, "Sparse semidefinite programs with near-linear time complexity," in *2018 IEEE Conference on Decision and Control (CDC)*. IEEE, 2018, pp. 1624–1631.



Steven H. Low (F'08) is the F. J. Gilloon Professor of the Department of Computing & Mathematical Sciences and the Department of Electrical Engineering at Caltech. Before that, he was with AT&T Bell Laboratories, Murray Hill, NJ, and the University of Melbourne, Australia. He has held honorary/chaired professorship in Australia, China and Taiwan. He was a co-recipient of IEEE best paper awards and is a Fellow of both IEEE and ACM. He was known for pioneering a mathematical theory of Internet congestion control and semidefinite relaxations of optimal power flow problems in smart grid. He received his B.S. from Cornell and PhD from Berkeley, both in electrical engineering.



Nicholas D. Sidiropoulos (F'09) earned the Diploma in Electrical Engineering from Aristotle University of Thessaloniki, Greece, and M.S. and Ph.D. degrees in Electrical Engineering from the University of Maryland at College Park, in 1988, 1990 and 1992, respectively. He has served on the faculty of the University of Virginia, University of Minnesota, and the Technical University of Crete, Greece, prior to his current appointment as Louis T. Rader Professor and Chair of ECE at UVA. From 2015 to 2017 he was an ADC Chair Professor at the University of Minnesota. His research interests are in signal processing, communications, optimization, tensor decomposition, and factor analysis, with applications in machine learning and communications. He received the NSF/CAREER award in 1998, the IEEE Signal Processing Society (SPS) Best Paper Award in 2001, 2007, and 2011, served as IEEE SPS Distinguished Lecturer (2008-2009), and as Vice President - Membership of IEEE SPS (2017-2019). He received the 2010 IEEE Signal Processing Society Meritorious Service Award, and the 2013 Distinguished Alumni Award from the University of Maryland, Dept. of ECE. He is a Fellow of IEEE (2009) and a Fellow of EURASIP (2014).



Fengyu Zhou (S'16) received the B.E. degree in electronics engineering from Tsinghua University, Beijing, China, in 2016 and the M.S. degree in electrical engineering from Caltech, Pasadena, CA, USA, in 2018. He is currently pursuing the Ph.D. degree in the Department of Electrical Engineering at Caltech.

His research interests are on the modeling, optimization and control of cyber-physical systems.



Ahmed S. Zamzam (Member, IEEE) is a Researcher at the National Renewable Energy Laboratory, Golden, CO, where he is a member of the Power Systems Engineering Center. He obtained his PhD in Electrical Engineering from the University of Minnesota, Minneapolis, MN, in 2019. Previously, he earned his BSc at Cairo University in 2013. Ahmed received the Louis John Schnell Fellowship (2015), and the Doctoral Dissertation Fellowship (2018) from the University of Minnesota.

His research interests include machine learning and optimization of smart grids, large-scale complex energy systems, and grid data analytics.